`ensembleBMA`: An `R` Package for Probabilistic Forecasting using Ensembles and Bayesian Model Averaging *

Chris Fraley, Adrian E. Raftery
Tilmann Gneiting, J. McLean Sloughter

Technical Report No. 516

Department of Statistics
University of Washington
Box 354322
Seattle, WA 98195-4322 USA

August 15, 2007

Abstract

`ensembleBMA` is a contributed `R` package for probabilistic forecasting using ensemble post-processing via Bayesian Model Averaging. It provides functions for parameter estimation via the EM algorithm for normal mixture models (appropriate for temperature or pressure) and mixtures of gamma distributions with a point mass at 0 (appropriate for precipitation) from training data. Also included are functions giving quantile forecasts based on these models, as well as for verification.

*Thanks go to Veronica Berrocal and Patrick Tewson for lending their expertise on a number of important issues, and to Michael Polakowski for his work on an earlier version of the package. We are also indebted to Cliff Mass, Jeff Baars, and Eric Grimit for many helpful discussions and for sharing data. Supported by the DoD Multidisciplinary Research Initiative (MURI) administered by the Office of Naval Research under grant N00014-01-10745 and by the Joint Ensemble Forecast System (JEFS) under University Corporation for Atmospheric Research subcontract S06-47225.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **15 AUG 2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE
ensembleBMA: An R Package for Probabilistic Forecasting using Ensembles and Bayesian Model Averaging

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Washington,Department of Statistics,Box 354322,Seattle,WA,98195-4322

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
ensembleBMA is a contributed R package for probabilistic forecasting using ensemble postprocessing via Bayesian Model Averaging. It provides functions for parameter estimation via the EM algorithm for normal mixture models (appropriate for temperature or pressure) and mixtures of gamma distributions with a point mass at 0 (appropriate for precipitation) from training data. Also included are functions giving quantile forecasts based on these models, as well as for verification.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**13**	

Contents

List of Tables

List of Figures

1 Overview

This document describes the ensembleBMA package for probabilistic forecasting using ensemble postprocessing via Bayesian Model Averaging, written in the R language. This package offers the following capabilities:

- Fitting BMA models to ensemble forecasting data with verification observations. Modeling capabilities consist of mixtures of normals for temperature and pressure and mixtures of gammas with a point mass at 0 for precipitation. The modeling can account for exchangeable ensemble members.

- Producing quantile forecasts from fitted BMA models.

- Computing continuous ranked probability scores and Brier scores for assessment of BMA forecasting performance.

- Displaying forecast and verification results.

An overview of the modeling methodology used in ensembleBMA can be found in Gneiting and Raftery (2005). More detail on the models and verification procedures can be found in Raftery et al. (2005), Sloughter et al. (2007), Gneiting et al. (2005), and Gneiting and Raftery (2007).

To use the ensembleBMA package, download it from the Comprehensive R Archive Network (CRAN) http://cran.r-project.org. Follow the instructions for installing R packages on your machine, and then do

```
> library(ensembleBMA)
```

inside R in order to use the software. Throughout this document it will be assumed that these steps have been taken before running the examples.

2 ensembleData objects

Modeling and forecasting functions in the ensembleBMA package require that the data be organized into an ensembleData object that includes the ensemble forecasts, usually with dates. Observed weather conditions are also needed for modeling and verification; other attributes such as latitude and longitude may be useful for plotting or analysis. The ensembleData object facilitates preservation of the data as a unit.

As an example, we create an ensembleData object called srftData corresponding to the srft data set of surface temperatures.

```
> data(srft)

> memberLabels <- c("CMCG","ETA","GASP","GFS","JMA","NGPS","TCWB","UKMO")

> srftData <- ensembleData(forecasts = srft[,memberLabels],
                 dates = srft$date, observations = srft$obs,
                 latitude = srft$lat, longitude = srft$lon)
```

It is advisable to assign member labels when creating `ensembleData` objects unless the matrix of forecasts labels the members, because they are used to match member names in data with the BMA model weights and parameters for forecasting and verification.

Specifying exchangeable ensemble members. Forecast ensembles may contain members that can be considered exchangeable or interchangeable; that is, their forecasts can be assumed to come from the same distribution. In such cases, parameters in the BMA model (including weights and bias correction coefficients) should be constrained to be the same among exchangeable members. In `ensembleBMA`, exchangeability can be specified when creating `ensembleData` objects by supplying a vector specifying a grouping of the ensemble members in the `exchangeable` argument. The non-interchangeable groups consist of singleton members, while exchangeable members would belong to the same group. As an illustration, suppose the `ETA` and `GFS` members are exchangeable in the example above, but all other members are non-interchangeable. The corresponding `ensembleData` object could be created as follows:

```
> data(srft)

> memberLabels <- c("CMCG","ETA","GASP","GFS","JMA","NGPS","TCWB","UKMO")

> exGroups <- c( CMCG=1, ETA=2, GASP=3, GFS=2, JMA=4, NGPS=5, TCWB=6, UKMO=7)

> srftDataX <- ensembleData(forecasts = srft[,memberLabels],
                  dates = srft$date, observations = srft$obs,
                  latitude = srft$lat, longitude = srft$lon,
                  exchangeable = exGroups)
```

The weights and parameters in the BMA model fit to `srtfDataX` will be equal for the `ETA` and `GFS` members.

3 BMA Forecasting

The `ensembleBMA` package provides several ways to obtain a forecast.

Surface Temperature Example. As an example, we model surface temperature for January 29, 2004 from ensemble forecasts and observations at station locations as given in the `srft` data set provided in the `ensembleBMA` package. The model fits a mixture of normals to the ensemble forecasts and observed data. We use the `srftData` object created in the previous section in the modeling. A training period of 25 days is used, with a lag of 2 days since the `srft` dataset gives 48-hour ensemble forecasts (Berrocal et al. 2007). The data is fitted with a mixture of normals as appropriate for temperature.

There are several options for obtaining the model. One is to use the function `ensembleBMA` with the date (or dates) of interest as input to obtain the associated BMA model (or models).

```
> srftBMA290104 <- ensembleBMA( srftData, dates = "2004012900",
          model = "normal", trainingRule = list(length = 25, lag = 2))
```

It should be noted that the `ensembleBMA` function will produce a model for any dates specified, provided that the dates and training rule are consistent with the available data. When no date is specified, the `ensembleBMA` function will produce a model for each date in the input data for which there is sufficient training data. The result of applying `ensembleBMA` with multiple dates can be used for forecasting on those dates.

The modeling process for a single date can also be separated into two steps: extraction of the training data for the desired date, and then fitting the model directly with `fitBMA`.

```
> train290104 <- trainingData( srftData, date = "2004012900",
                               trainingRule = list(length = 25, lag = 2))
> srftBMA290104fit <- fitBMA( train290104,  model = "normal")
```

A limitation of the two-step process is that date information is not retained as part of the model.

Forecasting is typically done on grids covering an area of interest rather than at station locations. The dataset `srftGrid` included in the `ensembleBMA` package gives forecasts of surface temperature initialized on January 27, 2004 and valid for January 29, 2004 at grid locations in the region in which the `srft` stations are located.

BMA forecasts for the grid locations can be obtained with `quantileForecastBMA`:

```
> data(srftGrid)

> memberLabels <- c("CMCG","ETA","GASP","GFS","JMA","NGPS","TCWB","UKMO")
> srftGridData <- ensembleData(forecasts = srftGrid[,memberLabels],
      latitude = srftGrid[,"latitude"], longitude = srftGrid[,"longitude"])
> gridForc290104 <- quantileForecastBMA( srftBMA290104, srftGridData,
                                        quantiles = c( .1, .5, .9))
```

The probability of freezing at grid locations can also be estimated using `cdfBMA`, which evaluates the cumulative distribution function for the model.

```
> probFreeze290104 <- cdfBMA( srftBMA290104, srftGridData, date = "2004012900",
                              value = 273.15)
```

In datasets `srft` and `srftGrid`, temperature is recorded in kelvins (K) corresponding to a freezing temperature of 273.15. The results can be displayed using the `plotBMAforecast` function, as shown below. Loading the `fields` library enables display of the country and state outlines, as well as a legend. A blue scale is chosen to display the probability of freezing, with darker shades representing higher probabilities.

```
> library(fields)

> plotBMAforecast( gridForc290104[,"0.5"], lon=srftGridData$lon,
                   lat=srftGridData$lat, type="image",
                   col=rev(rainbow(100,start=0,end=0.85)))
> title("Median Forecast for Surface Temperature", cex = 0.5)
```

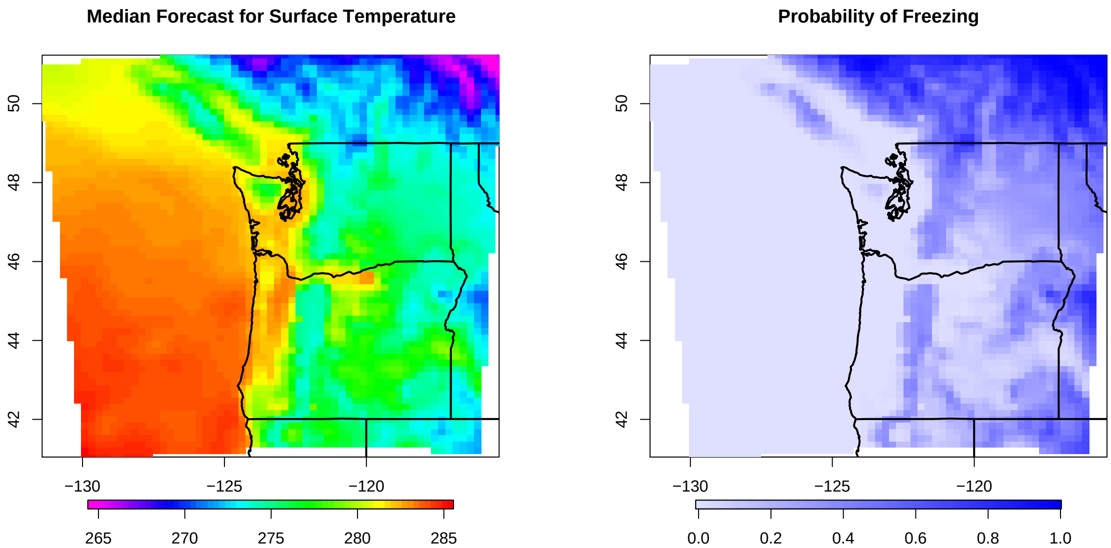

Figure 1: Image plots of the median BMA forecast of surface temperature and probability of freezing for January 29, 2004 from the `srftGrid` dataset. The plots were obtained by binning the forecasts at the grid locations onto a plotting grid. The `fields` library is used to allow addition of the legend and map outline to the plot.

```
> bluescale <- function(n)
               hsv(4/6, s = seq(from = 1/8, to = 1, length = n), v = 1)
> plotBMAforecast( probFreeze290104,
                   lon=srftGridData$lon, lat=srftGridData$lat,
                   type="image", col=bluescale(100))
> title("Probability of Freezing", cex = 0.5)
```

The resulting image plots are shown in Figure 1. The plots are made by binning values onto a plotting grid. The default (shown here) is to use binning rather than interpolation to determine these values.

Precipitation Example. In this example, we make use of the `prcpFit` and `prcpGrid` datasets included in the `ensembleBMA` package. The `prcpFit` dataset consists of the default BMA modeling parameters for the daily 48 hour forecasts of 24 hour accumulated precipitation (quantized to hundredths of an inch) over the US Pacific Northwest region from December 12, 2002 through March 31, 2005 used in Sloughter et al. 2007. The model fits a mixture of gamma distributions with a point mass at zero to the cube root transoformation of the ensemble forecasts and observed data. In this case the default training period of 30 days was used. The `prcpGrid` dataset consists of a grid of precipitation forecasts in the region of the observations used for `prcpFit` initialized on January 11, 2003 and valid for January 13, 2003.

```
> data(prcpGrid)
```

```
> prcpGridData <- ensembleData(forecasts = prcpGrid[,1:9],
```

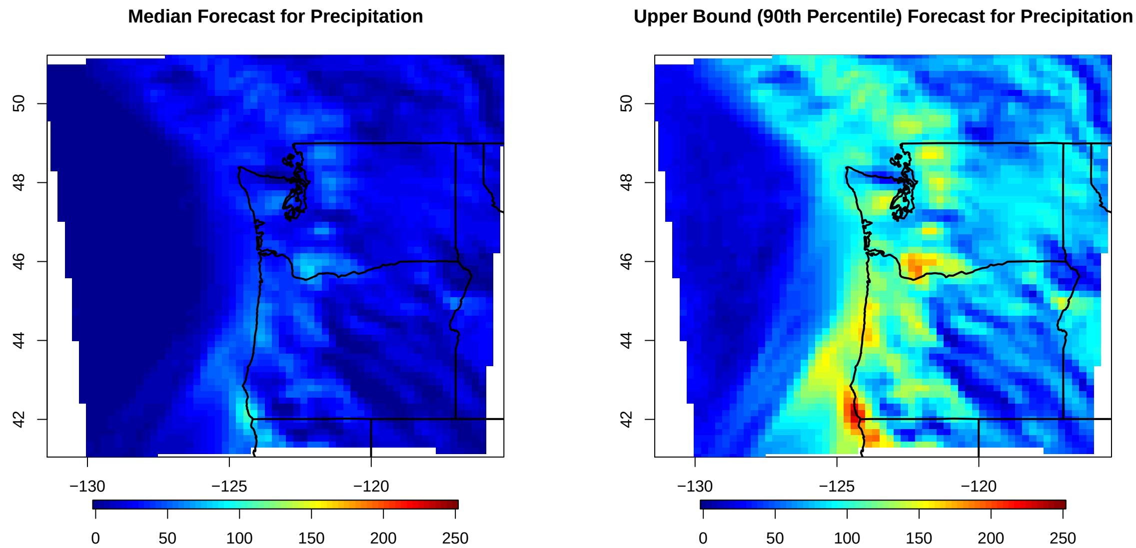

Figure 2: Image plots of the median and upper bound (90th percentile) BMA forecast of precipitation (measured in hundredths of an inch) for January 13, 2003 from the `prcpGrid` dataset. The plots were obtained by binning the forecasts at the grid locations onto a plotting grid. The `fields` library is used to allow addition of the legend and map outline to the plot.

```
                              latitude = prcpGrid[,"latitude"],
                              longitude = prcpGrid[,"longitude"])
```

The median and upper bound (90th percentile) forecasts can be obtained and plotted as follows:

```
> data(prcpFit)

> gridForc130103 <- quantileForecastBMA( prcpFit, prcpGridData,
                                          date = "20030113", q = c(0.5, 0.9))
> max(gridForc130103) # used to determine zlim in plotting
[1] 246.4196

> library(fields)

> plotBMAforecast( gridForc130103[,"0.5"], type = "image",
    zlim = c(0,250), lon=prcpGridData$lon, lat=prcpGridData$lat)
> title("Median Forecast for Precipitation", cex = 0.5)

> plotBMAforecast( gridForc130103[,"0.9"], type = "image",
    zlim = c(0,250), lon=prcpGridData$lon, lat=prcpGridData$lat)
> title("Upper Bound (90th Percentile) Forecast for Precipitation", cex = 0.5)
```

The corresponding plots are shown in Figure 2: The probability of precipitation and probability of precipitation above .25 inches can be obtained and plotted as follows. This gives an example of grayscale plotting of the data:

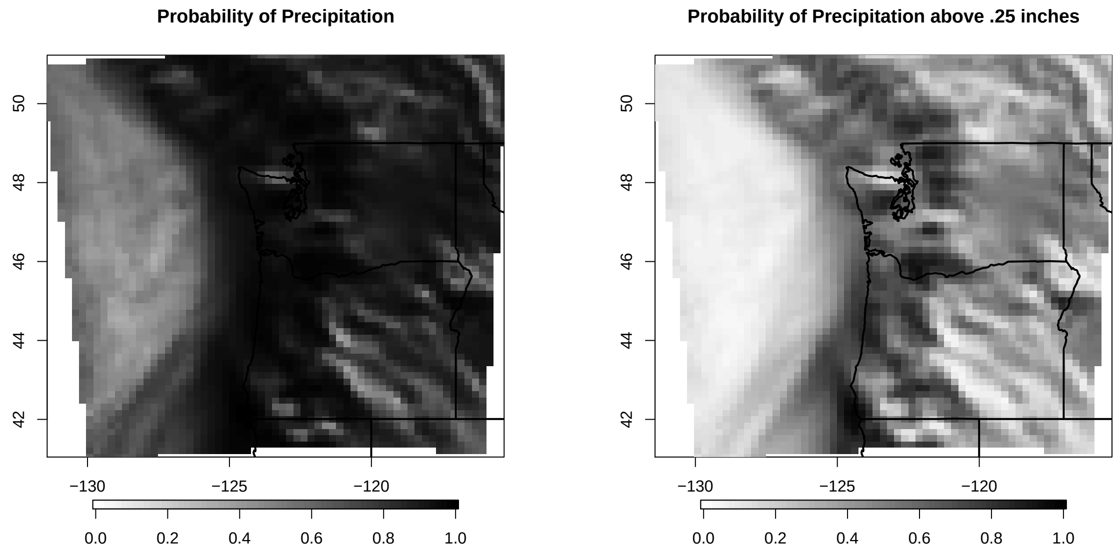

Figure 3: Grayscale image plots showing probability of precipitation for January 13, 2003 from the `prcpGrid` dataset. The plots were obtained by binning the forecasts at the grid locations onto a plotting grid. The `fields` library is used to allow addition of the legend and map outline to the plot.

```
> probPrecip130103 <- 1 - cdfBMA( prcpFit, prcpGridData, date = "20030113",
                                        values = c(0, 25))

> library(fields)

> grayscale <- function(n) gray((0:n)/n)

> range(probPrecip130103) # used to determine zlim in plots
[1] 0.02832709 0.99534860

> plotBMAforecast( probPrecip130103[,"0"],
                   lon=prcpGridData$lon, lat=prcpGridData$lat,
                   type="image", col= rev(grayscale(100)), zlim = c(0,1))
> title("Probability of Precipitation", cex = 0.5)

> plotBMAforecast( probPrecip130103[,"25"],
                   lon=prcpGridData$lon, lat=prcpGridData$lat,
                   type="image", col=rev(grayscale(100)), zlim = c(0,1))
> title("Probability of Precipitation above .25 inches", cex = 0.5)
```

The corresponding plots are shown in Figure 3.

4 Verification

The `ensembleBMA` package also provides a number of functions for verification. These can be applied to any data for which both a BMA forecasting model and observed weather conditions are available. Included are functions to compute mean absolute error, continuous ranked probability scores, and Brier scores.

Surface Temperature Example. In the previous section, we obtained a forecast of surface temperature on a grid of locations for January 29, 2004 from BMA modeling of station forecasts and observations from the `srft` data set provided in the `ensembleBMA` package. Forecasts can be obtained at the station locations by applying `quantileForecastBMA` to the model fit `srftBMA200104` from the previous section to the data used to generate the model.

```
> srftForc290104 <- quantileForecastBMA( srftBMA290104, srftData,
                                          quantiles = c( .1, .5, .9))
```

These forecasts can be plotted using `plotBMAforecast`. The example below shows contour plots in which the `R` core function `loess` has been used to interpolate the results at the station locations onto a grid for surface plotting.

```
> obs <- srftData$date == "2004012900"
> lat <- srftData$latitude[obs]; lon <- srftData$longitude[obs]

>  range(srftForc290104[,"0.5"]) # used to determine contour levels
[1] 265.1425 282.0040

> plotBMAforecast( srftForc290104[,"0.5"], lon, lat, interpolate = TRUE,
                   type = "contour", levels = seq(from=264, to=284, by=2))
> title("Median Forecast")
> points(lon, lat, pch = 16, cex = 0.5)  # observation locations

> plotBMAforecast( srftData$obs[obs], lon, lat, interpolate = TRUE,
                   type = "contour", levels = seq(from=264, to=284, by=2))
> title("Observed Surface Temperature")
> points(lon, lat, pch = 16, cex = 0.5)
```

The resulting plot is shown in Figure 4. In this case interpolation was used because the station locations are too sparse for binning. It is also possible to specify image or perspective plots, as well as contour plots. If the `fields` library is loaded, image plots will be enhanced as shown in the displays of the previous section.

The continuous ranked probability score (CRPS) and mean absolute error (MAE) (see, e.g. Gneiting and Raftery (2007)) can be obtained via functions `crps` and `mae`:

```
> crps( srftBMA290104, srftData)
ensemble      BMA
1.945544 1.490725

> mae( srftBMA290104, srftData)
ensemble      BMA
2.152070 2.042045
```

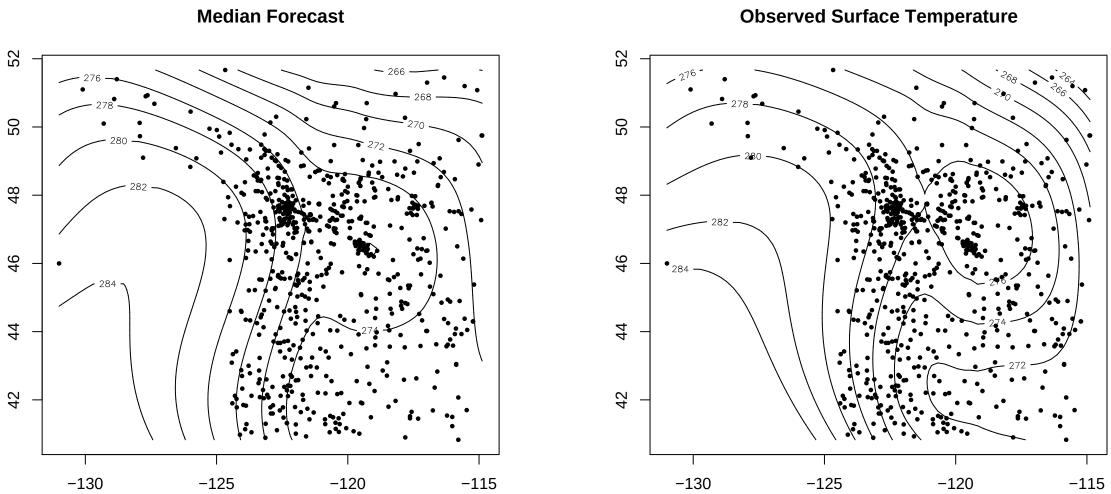

Figure 4: Contour plots of the BMA forecasts of surface temperature and verification observations at station locations for January 29, 2004 for the `srft` dataset. The plots were obtained using a `loess` fit to the forecasts and observations at the stations, interpolated on a plotting grid. The dots represent the 715 observation locations.

Here we are evaluating these measures for modeling at a single date; however, the CRPS and MAE would more typically be assessed over a range of dates and the corresponding models. For BMA mixtures of normals, `mae` computes the mean absolute difference of the BMA predictive mean (Raftery et al. 2005) and the observations.

Precipitation Example. In the previous section, we obtained a forecast of precipitation on a grid of locations for January 13, 2003 from BMA modeling of station forecasts and observations from the `prcpFit` and `prcpGrid` datasets provided in the `ensembleBMA` package Quantile forecasts can be obtained at the station locations by applying `quantileForecastBMA` to the model fit given the data used to generate the model. An `ensembleBMAobject` called `prcpDJdata` is provided as a dataset with the package containing ensemble forecasts and verification observations for this date.

```
> prcp130103 <- quantileForecastBMA( prcpFit, prcpDJdata, date = "20030113",
                                     quantiles = c( .1, .5, .9))
```

We can compare the forecasts with the observed data graphically as follows:

```
> obs <- prcpDJdata$date == "20030113"
> lon <- prcpDJdata$longitude[obs]; lat <- prcpDJata$latitude[obs]
> verif <- prcpDJdata$obs[obs]

> nObs <- sum(obs)
> forc10 <- prcp130103[,"0.1"]
> forc50 <- prcp130103[,"0.5"]
```

```
> forc90 <- prcp130103[,"0.9"]
> ord <- order(forc90)
> ylim <- c(0, max(forc90,verif))
> plot(1:nObs, forc90[ord], ylim = ylim, type = "l", col = "black",
     xlab = "", ylab = "", xaxt = "n")
> lines(1:nObs, forc10[ord], type = "l", col = "gray")
> lines(1:nObs, forc50[ord], type = "l", col = "red")
> points(1:nObs, verif, pch = 16, col = "black", cex = 0.5)
> title("Forecasts and Observations for January 13, 2003", cex = 0.5)
```

The resulting plots are shown in Figure 5.

The continuous ranked probability score (CRPS) and mean absolute error (MAE) can be obtained via functions `crps` and `mae`. Here we have done so for the entire precipitation data set available from `http://www.stat.washington.edu/MURI`. The object `prcpData` is the `ensembleData` object of that data set used to obtain the models in `prcpFit`. It is not included in the `ensembleBMA` package on account of its size.

```
> crps( prcpFit, prcpData)
 ensemble      BMA
 7.545675 5.597090

> mae(prcpFit, prcpData)
ensemble      BMA
9.924270 7.484926
```

For BMA mixtures of gammas with a point mass at 0, `mae` computes the mean absolute difference of the BMA median forecast and the observations (Sloughter et al. 2007). Brier scores (see, e.g. Joliffe and Stephenson, 2003) for the model fits can be obtained via the function `brierScore`.

```
>  brierScore( prcpFit, prcpData, thresh = c(0, 50, 100, 200, 300, 400))
  thresholds  climatology     ensemble     logistic          bma
1          0 0.2238453660 0.2685776155 0.1419662590 0.1409636402
2         50 0.0436537385 0.0433243610 0.0299279717 0.0321579470
3        100 0.0165649433 0.0153095694 0.0121539204 0.0131247108
4        200 0.0041069368 0.0036680461 0.0036204314 0.0035433308
5        300 0.0016872843 0.0015887788 0.0016012870 0.0015381613
6        400 0.0008256083 0.0008142352 0.0008194794 0.0007742297
```

Here ‘climatology’ refers to the empirical distribution of the verifying observations, while ‘logistic’ refers a logistic regression model with the cube root of the data as predictor variable, with coefficients determined from the training data. This logistic regression model is the one used for the probability of precipitation component in the forecasting model of Sloughter et al. (2007).

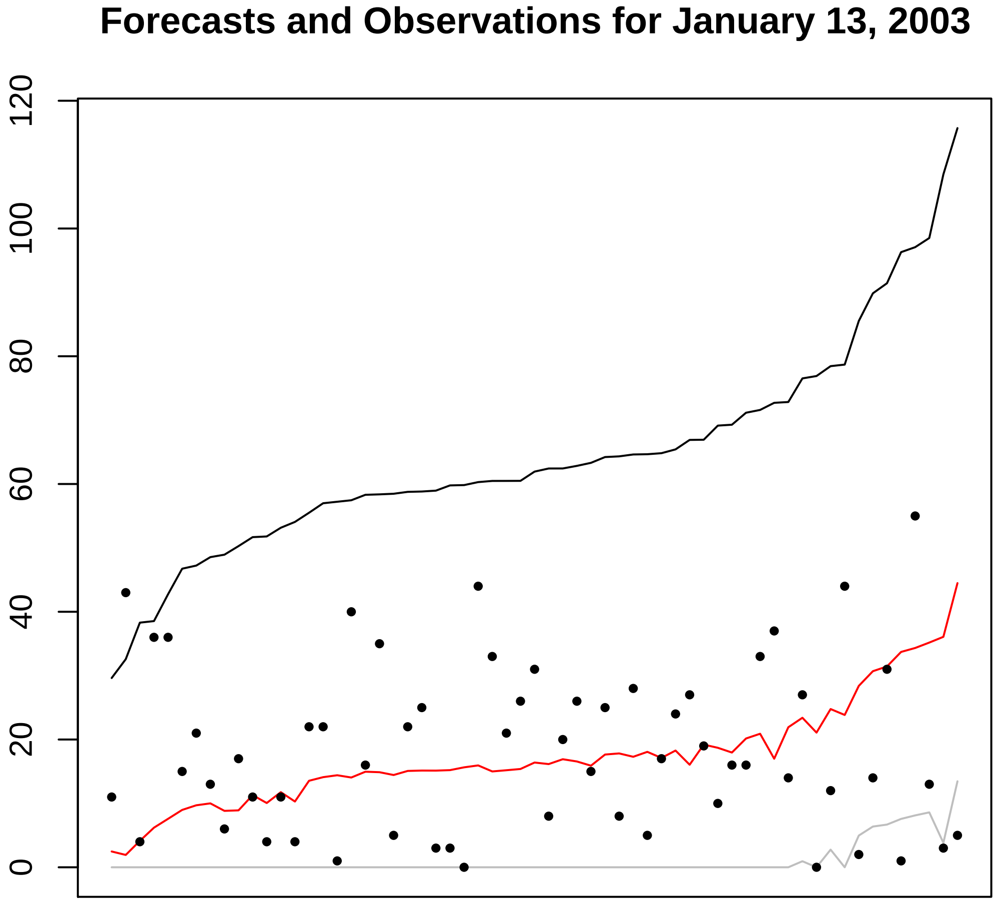

Figure 5: The lines represent the 10th (gray), 50th (red), and 90th (black) percentile BMA forecasts of precipitation for January 13, 2003 at the station locations, while the dots indicate the observed precipitation at the same locations. The horizontal axis represents the observations, in order of increasing 90th precentile forecast.

References

[1] V. J. Berrocal, A. E. Raftery, and T. Gneiting. Combining spatial statistical and ensemble information in probabilistic weather forecasts. *Monthly Weather Review*, 135:1386–1402, 2007.

[2] V. J. Berrocal, A. E. Raftery, T. Gneiting, and R. C. Steed. Probabilistic weather forecasting for winter road maintenance. Technical Report 511, University of Washington, April 2007.

[3] T. Gneiting and A. E. Raftery. Weather forecasting with ensemble methods. *Science*, 310:248–249, 2005.

[4] T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378, 2007.

[5] T. Gneiting, A. E. Raftery, A. Westveld, and T. Goldman. Calibrated probabilitistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review*, 133:1098–1118, 2005.

[6] I. T. Jolliffe and D. B. Stephenson, editors. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Wiley, 2003.

[7] A. E. Raftery, T. Gneiting, F. Balabdaoui, and M. Polakowski. Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133:1155–1174, 2005.

[8] J. M. Sloughter, A. E. Raftery, T. Gneiting, and C. Fraley. Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Monthly Weather Review*, 135:3209–3220, 2007.